3/9/1
DIALOG(R)File 155:MEDLINE(R)

13263977 22016913 PMID: 12022223
Mutation Master: profiles of substitutions in hepatitis C virus RNA of the core, alternate reading frame, and NS2 coding regions.
Walewski Jose L; Gutierrez Julio A; Branch-Elliman Westyn; Stump Decherd D; Keller Toby R; Rodriguez Alfredo; Benson Gary; Branch Andrea D
Department of Medicine, Division of Liver and Diseases, Mount Sinai School of Medicine, New York, New York 10029, USA.
RNA (New York, N.Y.) (United States) May 2002, 8 (5) p557-71, ISSN 1355-8382 Journal Code: 9509184
Contract/Grant No.: R01 DK52071; DK; NIDDK
Document type: Journal Article
Languages: ENGLISH
Main Citation Owner: NLM
Record type: Completed
Subfile: INDEX MEDICUS
The RNA genome of the hepatitis C virus (HCV) undergoes rapid evolutionary change. Efforts to control this virus would benefit from the advent of facile methods to identify characteristic features of HCV RNA and proteins, and to condense the vast amount of mutational data into a readily interpretable form. Many HCV sequences are available in GenBank. To facilitate analysis, consensus sequences were constructed to eliminate the overrepresentation of certain genotypes, such as genotype 1, and a novel package of sequence analysis tools was developed. Mutation Master generates profiles of point mutations in a population of sequences and produces a set of visual displays and tables indicating the number, frequency, and character of substitutions. It can be used to analyze hundreds of sequences at a time. When applied to 255 HCV core protein sequences, Mutation Master identified variable domains and a series of mutations meriting further investigation. It flagged position 4, for example, where 90% or more of all sequences in genotypes 1, 2, 4, and 5, have N4, whereas those in genotypes 3, 6, 7, 8, 9, and 10 have L4. This pattern is noteworthy: L (hydrophobic) to N (polar) substitutions are generally rare, and genotypes 1, 2, 4, and 5 do not form a recognized super family of sequences. Thus, the L4N substitution probably arose independently several times. Moreover, not one member of genotypes 1, 2, 4, or 5 has L4 and not one member of genotypes 3, 6, 7, 8, 9, or 10 has N4. This nonoverlapping pattern suggests that coordinated changes at position 4 and a second site are required to yield a viable virus. The package generated a table of genotype-specific substitutions whose future analysis may help to identify interacting amino acids. Three substitutions were present in 100% of genotype 2 members and absent from all others: A68D, R74K, and R114H. Finally, this study revealed thatARFP, a novel protein encoded in an overlapping reading frame, is as conserved as conventional HCV proteins, a result supporting a role for ARFP in the viral life cycle. Whereas most conventional programs for phylogenetic analysis of sequences provide information about overall relatedness of genes or genomes, this program highlights and profiles point mutations. This is important because determinants of pathogenicity and drug susceptibility are likely to result from changes at only one or two key nucleotides or amino acid sites, and would not be detected by the type of pairwise comparisons that have usually been performed on HCV to date. This study is the first application of Mutation Master, which is now available upon request (http://tandem.biomath.mssm.edu/mutationmaster.html).
Tags: Human; Support, U.S. Gov't, Non-P.H.S.; Support, U.S. Gov't, P.H.S.
Descriptors: *Hepacivirus--genetics--GE; *Mutation; *RNA, Viral--genetics --GE; *Software; Amino Acid Motifs; Amino Acid Sequence; Computational Biology; Consensus Sequence; Genome, Viral; Genotype; Molecular Sequence Data; Point Mutation; Protein Structure, Tertiary; RNA, Viral--chemistry --CH; Reading Frames; Sequence Alignment--statistics and numerical data--SN ; Viral Core Proteins--genetics--GE; Viral Nonstructural Proteins--genetics --GE
Molecular Sequence Databank No.: GENBANK/AF011751

CAS Registry No.: 0 (NS2 protein, hepatitis C virus); 0 (RNA, Viral); 0 (Viral Core Proteins); 0 (Viral Nonstructural Proteins)
Record Date Created: 20020522

3/9/3
DIALOG(R)File 155:MEDLINE(R)

11218139 21247671 PMID: 11350035
Evidence for a new hepatitis C virus antigen encoded in an overlapping reading frame.
Walewski J L; Keller T R; Stump D D; Branch A D
Department of Medicine, Mount Sinai School of Medicine, New York, New York 10029, USA.
RNA (New York, N.Y.) (United States) May 2001, 7 (5) p710-21, ISSN 1355-8382 Journal Code: 9509184
Contract/Grant No.: P01 DK50795; DK; NIDDK; R01 DK52071; DK; NIDDK
Document type: Journal Article
Languages: ENGLISH
Main Citation Owner: NLM
Record type: Completed
Subfile: INDEX MEDICUS
Many viruses have overlapping genes and/or regions in which a nucleic acid signal is embedded in a coding sequence. To search for dual-use regions in the hepatitis C virus (HCV), we developed a facile computer-based sequence analysis method to map dual-use regions in coding sequences. Eight diverse full-length HCV RNA and polyprotein sequences were aligned and analyzed. A cluster of unusually conserved synonymous codons was found in the core-encoding region, indicating a potential overlapping open reading frame (ORF). Four peptides (A1, A2, A3, and A4) representing this alternate reading frame protein (ARFP), two others from the HCV core protein, and one from bovine serum albumin (BSA) were conjugated to BSA and used in western blots to test sera for specific antibodies from 100 chronic HCV patients, 44 healthy controls, and 60 patients with non-HCV liver disease. At a 1:20,000 dilution, specific IgGs to three of the four ARFP peptides were detected in chronic HCV sera. Reactivity to either the A1 or A3 peptides (both ARFP derived) was significantly associated with chronic HCV infection, when compared to non-HCV liver disease serum samples (10/100 versus 1/60; $p < 0.025$). Antibodies to A4 were not detected in any serum sample. Our western blot assays confirmed the presence of specific antibodies to a new HCV antigen encoded, at least in part, in an alternate reading frame (ARF) overlapping the core-encoding region. Because this novel HCV protein stimulates specific immune responses, it has potential value in diagnostic tests and as a component of vaccines. This protein is predicted to be highly basic and may play a role in HCV replication, pathogenesis, and carcinogenesis.
Tags: Human; Support, Non-U.S. Gov't; Support, U.S. Gov't, P.H.S.
Descriptors: *Alternative Splicing; *Antigens, Viral--genetics--GE; *Conserved Sequence; *Genes, Overlapping; *Hepacivirus--genetics--GE; *Open Reading Frames; *Viral Core Proteins--genetics--GE; Amino Acid Sequence; Antigens, Viral--immunology--IM; Binding, Competitive; Blotting, Western; Genes, Viral; Hepacivirus--immunology--IM; Hepatitis C Antibodies--blood --BL; Hepatitis C Antibodies--immunology--IM; Hepatitis C, Chronic--blood --BL; Hepatitis C, Chronic--immunology--IM; Hepatitis C, Chronic --virology--VI; Molecular Sequence Data; Peptides--genetics--GE; Peptides --immunology--IM; Sensitivity and Specificity
CAS Registry No.: 0 (Antigens, Viral); 0 (Hepatitis C Antibodies); 0 (Peptides); 0 (Viral Core Proteins); 0 (hepatitis C virus nucleocapsid protein)
Record Date Created: 20010514
?

1299624 21340276 PMID: 11447125

Synthesis of a novel hepatitis C virus protein by ribosomal frameshift.

Xu Z; Choi J; Yen T S; Lu W; Strohecker A; Govindarajan S; Chien D; Selby M J; Ou J

Department of Molecular Microbiology and Immunology, University of Southern California, 2011 Zonal Avenue, HMR-401, Los Angeles, CA 90033, USA.

EMBO journal (England) Jul 16 2001, 20 (14) p3840-8, ISSN 0261-4189 Journal Code: 8208664

Contract/Grant No.: R03AI45873; AI; NIAID; U019AI40038; AI; NIAID

Document type: Journal Article

Languages: ENGLISH

Main Citation Owner: NLM

Record type: Completed

Subfile: INDEX MEDICUS

Hepatitis C virus (HCV) is an important human pathogen that affects approximately 100 million people worldwide. Its RNA genome codes for a polyprotein, which is cleaved by viral and cellular proteases to produce at least 10 mature viral protein products. We report here the discovery of a novel HCV protein synthesized by ribosomal frameshift. This protein, which we named the F protein, is synthesized from the initiation codon of the polyprotein sequence followed by ribosomal frameshift into the -2/+1 reading frame. This ribosomal frameshift requires only codons 8-14 of the core protein-coding sequence, and the shift junction is located at or near codon 11. An F protein analog synthesized in vitro reacted with the sera of HCV patients but not with the sera of hepatitis B patients, indicating the expression of the F protein during natural HCV infection. This unexpected finding may open new avenues for the development of anti-HCV drugs.

Tags: Support, Non-U.S. Gov't; Support, U.S. Gov't, P.H.S.

Descriptors: *Frameshifting, Ribosomal; *Hepacivirus--metabolism--ME; *Viral Core Proteins--genetics--GE; Amino Acid Sequence; Base Sequence; Codon; DNA, Viral; Genome, Viral; Hepacivirus--genetics--GE; Molecular Sequence Data; Open Reading Frames; Viral Core Proteins--chemistry--CH

CAS Registry No.: 0 (Codon); 0 (DNA, Viral); 0 (Viral Core Proteins); 0 (hepatitis C protein F)

```
? ds

Set      Items    Description
S1      268085    (HEPATITIS (W) C)
S2         264    (ALTERNATE (W) READING (W) FRAME?)
S3         735    (OVERLAPPING (W) READING (W) FRAME?)
S4       54201    (FRAME (W) SHIFT?) OR FRAMESHIFT?
S5       55046    S2 OR S3 OR S4
S6         603    S1 AND S5
S7         417    RD (unique items)
S9     3325540    CORE OR CAPSID
S10        167    S7 AND S9
```